

# **Why People Cross Where They Do**

## **The Role of the Street Environment**

TRB Paper No. 03-3078

Xuehao Chu  
Center for Urban Transportation Research  
University of South Florida  
4202 East Fowler Avenue, CUT 100  
Tampa, Florida 33620  
xchu@cutr.usf.edu  
Tel.: (813) 974-9831  
Fax: (813) 974-5168

Martin Guttenplan  
Florida Department of Transportation  
Systems Planning Office  
605 Suwannee St. MS-19  
Tallahassee, FL 32399  
martin.guttenplan@dot.state.fl.us  
Tel.: (850) 414-4906  
Fax: (850) 921-6361

Mike Baltes  
Center for Urban Transportation Research  
University of South Florida  
4202 East Fowler Avenue, CUT 100  
Tampa, Florida 33620  
baltes@cutr.usf.edu  
Tel.: (813) 974-9843  
Fax: (813) 974-5168

Revised  
September 2002

This paper models the role of the street environment in how people cross roads in urban settings. Respondents were placed in real traffic conditions at the curbside of street blocks in the Tampa Bay area for a three-minute observation of the street environment. Without crossing the blocks, each respondent stated his crossing preference at each of six blocks. The origin and destination of each crossing were hypothetically set and varied across the blocks. So were the options available: two options for crossing at an intersection and up to four options for crossing at mid-block locations. Within the framework of discrete-choice models, the stated preferences are explained with the street environment, including traffic conditions, roadway characteristics, and signal-control characteristics. All three components of the street environment are considered: mid-block locations, intersections, and the roadside environment. The paper describes survey design and data collection efforts; estimates a nested logit model of pedestrian street-crossing behavior; and discusses its implications to researchers and practitioners.

1. Report No. <b>NCTR-473-06</b>	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle <b>Why People Cross Where They Do: The Role of the Street Environment</b>		5. Report Date <b>September 2002</b>	
		6. Performing Organization Code	
7. Author(s) <b>Xuehao Chu, Martin Guttenplan, and Michael R. Baltes</b>		8. Performing Organization Report No.	
9. Performing Organization Name and Address <b>National Center for Transit Research (NCTR) University of South Florida 4202 E Fowler Av, CUT 100, Tampa, FL 33620-5375</b>		10. Work Unit No.	
		11. Contract or Grant No. <b>DTRS98-G-0032</b>	
12. Sponsoring Agency Name and Address <b>Office of Research and Special Programs U.S. Department of Transportation, Washington, D.C. 20590</b>  <b>Florida Department of Transportation 605 Suwannee Street, MS 26, Tallahassee, FL 32399</b>		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes <b>Supported by a grant from the Florida Department of Transportation and the U.S. Department of Transportation</b>			
16. Abstract  This paper models the role of the street environment in how people cross roads in urban settings. Respondents were placed in real traffic conditions at the curbside of street blocks in the Tampa Bay area for a three-minute observation of the street environment. Without crossing the blocks, each respondent stated his crossing preference at each of six blocks. The origin and destination of each crossing were hypothetically set and varied across the blocks. So were the options available: two options for crossing at an intersection and up to four options for crossing at mid-block locations. Within the framework of discrete-choice models, the stated preferences are explained with the street environment, including traffic conditions, roadway characteristics, and signal-control characteristics. All three components of the street environment are considered: mid-block locations, intersections, and the roadside environment. The paper describes survey design and data collection efforts; estimates a nested logit model of pedestrian street-crossing behavior; and discusses its implications to researchers and practitioners.			
17. Key Words <b>Pedestrian, street crossing, mid-block, intersection, roadside, nested logit, stated preference</b>	18. Distribution Statement <b>Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161, 703-487-465, and through the NCTR website at <a href="http://www.nctr.usf.edu">http://www.nctr.usf.edu</a></b>		
19. Security Classif. (of this report) <b>Unclassified</b>	20. Security Classif. (of this page) <b>Unclassified</b>	21. No. of pages <b>14 (without appendices)</b>	22. Price

## INTRODUCTION

Street crossing is a critical element of the urban transportation environment for pedestrians. A large body of work already exists on street crossing by pedestrians, including the following by subject area:

- Crossing delays (1),
- Crossing opportunities (2-3),
- Pedestrians' behavioral parameters such as walking speed, start-up time, and gap-acceptance (4-6),
- Pedestrian compliance (7),
- Pedestrian perceptions toward specific treatments (8-9),
- Determination of level of service (10-13),
- Engineering parameters such as pedestrian clearance intervals (14),
- Evaluation of treatments (15-18),
- Drivers' perspective, including pedestrian visibility, effect of crosswalk markings, non-compliance with signals (19-21),
- Safety (22-23), and
- Empirical modeling (24-26).

However, little research exists that can help answer questions related to pedestrian planning, engineering solutions to pedestrian crossing safety, and research methods for modeling street-crossing behavior. Below are a few examples of these questions:

### **Planning Questions**

- How can existing planning tools for determining pedestrian level of service for street crossing at mid-block locations and intersections be integrated to determine pedestrian level of service at the block level?

### **Engineering Questions**

- How and when might a pedestrian go to a marked crosswalk in mid-block locations?
- How and when might a pedestrian go to an intersection for street crossing?
- Where should transit bus stops be located so that transit users are more likely to choose safe crossing options to access them?

### **Research Methodology Questions**

- What statistical models are most appropriate for modeling the street-crossing behavior of pedestrians so that these planning and engineering questions can be answered?
- What and how should data be collected in order to estimate such statistical models?

This paper models the role of the street environment in how people cross roads in urban settings. Specifically, 86 participants placed in real traffic conditions at the curbside of 48 street blocks in the Tampa Bay area observed the street environment for three minutes. Without crossing the street blocks, each participant stated his crossing preference at each of six blocks. The origin and destination for each crossing were hypothetically set and varied across the blocks. So were the options available: two options for crossing at an intersection and up to four options for crossing at a mid-block location. Within the framework of discrete choice models, the stated preferences are explained by traffic conditions, roadway characteristics, and signal-control characteristics.

The paper focuses on the street environment so that all variables can be readily measured for model applications. As an alternative, one could model the role of the direct attributes, such as safety and time, that pedestrians may tradeoff in choosing a crossing option. By focusing on the street environment, the paper assumes that the indirect attributes that characterize the street environment determine the direct attributes and that the street crossing behavior can be modeled with these indirect attributes equally well. As another alternative, one could include the street environment as well as pedestrians' personal characteristics. It is recognized here that these characteristics are potentially important in how people cross roads. They are excluded solely because data on them are not readily available for model applications. The impacts of these two alternative specifications on model results are reported elsewhere (27) and are briefly described in this paper when its research implications are discussed.

The rest of the paper has four sections. They describe: 1) the design of the stated-preference survey, 2) data collection efforts, 3) model estimation results, and 4) shortcomings of the study and its implications to pedestrian planning, engineering solutions to pedestrian safety for street crossing, and research, respectively.

## SURVEY DESIGN

The stated-preference approach was chosen for several reasons. It resulted in wide ranges of variation in the street environment. It allowed solicitation of crossing preferences in real traffic conditions. It also resulted in a manageable number of crossing options for modeling. The design process for this research involved four steps:

1. Identify potential determinants of pedestrian street-crossing behavior;
2. Determine levels of key determinants through the selection of street blocks;
3. Formulate crossing scenarios by defining crossing origins and destinations, crossing options, and temporary mid-block crosswalks; and
4. Develop instruments for individual crossing scenarios.

These reasons and design steps differ from those for a standard stated-preference survey (28).

### Potential Determinants

Two steps were used to select potential determinants that describe the street environment. The first step identified the direct attributes that pedestrians may tradeoff in making a choice: comfort, safety, time, and predictability. Predictability refers to the uncertainty in the amount of time an option may take a pedestrian to cross. The second step identified the indirect factors that may determine the direct attributes.

#### *Comfort and Predictability*

Differences in comfort result largely from differences in exposure to unpleasantness (such as hot weather) and personal traits that influence comfort sensitivity (such as poor health). Such differences are captured with roadside walking and crossing distance. Roadside walking could vary significantly across options. Crossing distance varies when jaywalking is involved or when the choice involves intersections and mid-block locations that have different width. Variation in predictability results from the presence or absence as well as the spacing of traffic signals.

#### *Safety and Time*

The amount of time spent walking along a street is determined by the distance involved and speed of walking. Distance is already identified as a potential factor in the paragraph above. The potential factors for safety, crossing time, and waiting time are discussed below for crossing at mid-block locations, crossing at intersections, and roadside walking separately.

**Mid-block Locations.** Chu and Baltes (29) identify potential determinants for pedestrian crossing behavior at mid-block locations, based on supply of gaps, crossing time, and safety margin, which form the three components of the gap-acceptance behavior of pedestrians (24). Safety margin is the difference between the time a pedestrian takes to cross the traffic and the time the next vehicle arrives at the crossing point.

**Intersections.** Crider et al. identify potential determinants for pedestrian crossing behavior at intersections (11). These are done separately for safety and delays. Safety consists of conflicts with motor vehicles and pedestrian's exposure to these conflicts. Vehicle movements at an intersection that cross the crosswalk represent conflict volumes. Exposure consists of crossing distance, presence of crosswalks, and presence of curb or sidewalk, and median type. For pedestrian delays, the potential determinants differ between signalized and un-signalized intersections. At signalized intersections, pedestrian crossing delay depends on cycle length for crossing with a pedestrian signal and on the facility's green ratio for crossing without a pedestrian signal. At un-signalized intersections, pedestrian crossing delay is a function of the conflict volumes described above.

**Roadside.** Landis et al. identify a set of potential determinants for pedestrians walking along roadsides (30). Through a step-wise regression process, the authors identify a number of factors describing the roadside environment, including the various components of lateral separation between sidewalks and traffic lanes.

### Site Selection

The selection of blocks for the field survey determined the values for most aspects of the street environment and the combinations of these values. The following criteria were used:

1. All blocks had two intersecting roads at the two ends with through movement.
2. All blocks were on roads that are functionally classified as collector or above in urban settings.

3. The blocks were from different regions of the Tampa Bay area. In order to facilitate survey logistics, the selection was further limited to a circle of 5-mile radius within each of four sub-areas: northeast Tampa, South Tampa, Clearwater, and St. Petersburg.
4. A number of potential determinants were considered, including number of lanes, presence and type of medians, signalization and crosswalk marking at intersections, pedestrian signal heads at intersections, sidewalks, lateral separation between sidewalks and traffic lanes, and block length.
5. A wide range of combinations of the values of the considered determinants was included. For example, it is desirable to have blocks on a 6-lane road with medians and blocks on a 6-lane road without medians.
6. A total of 48 blocks were selected with 12 from each area. The number 48 was chosen because it resulted in 12 blocks in each area. Field surveys were done on different days in the different areas. Furthermore, the 12 blocks in each area were divided into two groups of 6 each. These two 6-block groups were visited by two different groups of survey participants with each group taken by a bus. Based on the survey experience reported by Baltes and Chu (10), a single bus was able to visit six sites in a single day.

The actual selection was a manual process with hundreds of miles of driving and several steps:

- Produce GIS maps that show roads classified as collector or above within each circle.
- Identify blocks in the field that meet criterion 1 and record information on the determinants in criterion 4.
- Based on the information from the field, select 12 candidate blocks within each area that meet criterion 5.
- Check selected blocks in the field and adjust when needed.

### **Crossing Scenarios**

A crossing scenario is what was presented to a survey participant for soliciting his stated crossing preferences. a crossing scenario for a block consisted of the street environment, the origin and destination of the crossing, and the crossing options available to the pedestrian for the particular origin and destination. Much of the street environment for any block was determined once it was included in the sample of blocks. The only exception was crosswalk markings, particularly at mid-block locations. In addition to defining individual crossing scenarios, the design process determined what set of crossing scenarios each survey participant was presented with.

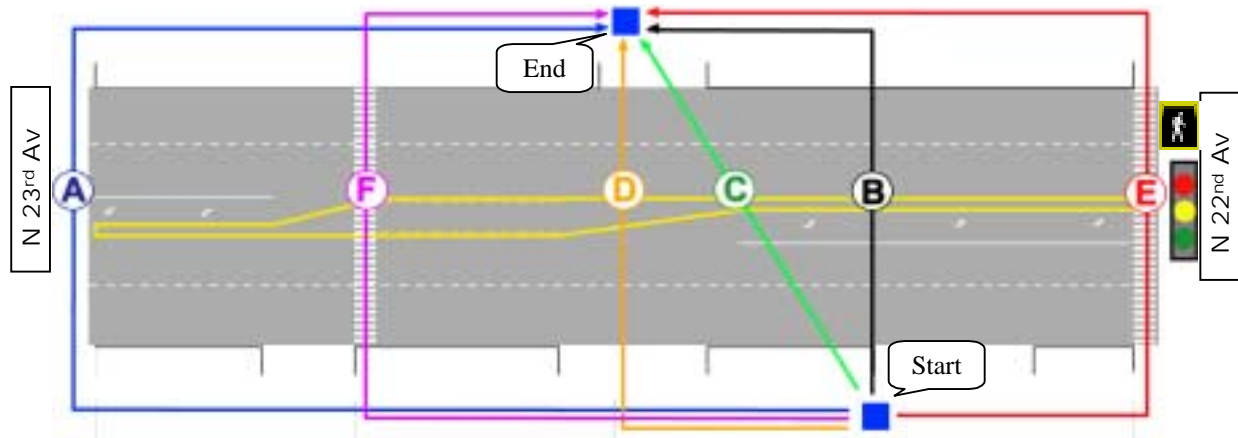
#### *Start and End Points*

The origin and destination for any crossing scenario were called the start and end points (Figure 1). Five potential locations for either the start or end point were considered with equal distance between them. For either the start or end point, two potential locations were at the intersections. These potential locations allowed a total of 25 different start-end combinations. Two combinations of start and end points were randomly selected for each block. For ease of reference, the side of a block with the start point was called the nearside and the other side the farside.

Figure 1. Sample Survey Instrument for Stated Preferences

Please enter your PIN here: \_\_\_\_\_

The diagram below shows your start point, your end point, and your location options for crossing the street within this block.



Please stand at your start point and observe the block characteristics and traffic conditions for 3 minutes. Based on your observation of the block and evaluation of the options during these 3 minutes, please tell us your choice for crossing this street by selecting **one** from below:

**A**

**F**

**D**

**C**

**B**

**E**

### *Crosswalk Marking*

Mid-block crosswalks rarely exist in the study area. In fact, none of the 48 street blocks had a mid-block crosswalk. Temporary marking was instead used to define mid-block crosswalks. About half of the sample blocks had a temporary mid-block crosswalk with three in each six-block group. A manual process was used to determine which three blocks in a six-block group got a mid-block crosswalk or where a mid-block crosswalk was placed on a given block. This determination was made visually with simultaneous consideration of all blocks in the same six-block group and with factors considered shown graphically. Factors considered include roadway width, block length (short, medium, long), presence and type of medians, crosswalk marking at intersections, traffic signals, pedestrian signals at intersections, and the two chosen start-end combinations.

Three materials for marking crosswalks were tested on two clear days, on two blocks, on a six-lane road, with 12 participants: pavement tapes, chalk powder, and four orange traffic cones with two on each side of the road. The question for the test participants was: Did the marking adequately represent a marked crosswalk to you during the test? The answers were on a 1-5 scale with 5 being adequate and 1 inadequate. Chalk powder was easily blown off by passing motor vehicles. Both orange cones and pavement tapes were perceived to be adequate to represent real crosswalk marking and orange traffic cones were as effective as pavement tapes. Orange traffic cones were chosen over pavement tapes for logistical, material cost, and safety reasons.

### *Crossing Options*

For a given start-end combination, a set of up to six discrete options was defined that can approximate most of the potentially infinite number of crossing options. These options are labeled as A through F for ease of reference and defined as follows (left and right are relative to the nearside):

- A = Crossing at the left intersection (left intersection)
- B = Crossing at a mid-block start point at a right angle (cross first and walk later)
- C = Crossing with a jaywalk between the start and end points (jaywalk)
- D = Walking to the opposite of a mid-block end point and crossing there at a right angle (walk first and cross later)
- E = Crossing at the right intersection (right intersection)
- F = Crossing at a mid-block crosswalk that is away from a start or end point (mid-block crosswalk)

The phrases in the parentheses may be used to refer to these options.

The exact options vary. The availability of options A through E depends on the particular start-end combination. If both the start and end points are located at mid-block locations but not across each other, for example, options A through E would all be available. If the start point is at the left intersection instead, option B would disappear and option A would no longer involve walking along the nearside. If the start and end points are located at the same intersection, only A and E would be available. In general, there are a total of five possible sets of options from the 25 possible start-end combinations discussed earlier. These are: A-E, A-C-E, A-B-E, A-C-D-E, A-B-C-E, and A-B-C-D-E. On the other hand, Option F is available only when a mid-block crosswalk is present and located away from a start or end point. All options are available in the diagram in Figure 1.

### *Group Crossing Scenarios*

Each survey participant provided 12 stated-preference responses with 2 responses on each of six blocks in the same circle. As discussed earlier, two combinations of start and end points were selected for each block. The two responses for a given block from the same participant were for these two different start-end combinations. More is discussed on how these two responses were obtained in the section on field surveys. The particular six blocks within the same area were determined with two considerations. The six blocks would result in a route that is similar in length with the other six blocks in the same circle. Each six-block group would have as much variation as possible in key determinants.

### **Instruments**

There were a total of 96 instruments with each block having two of them, corresponding to the two combinations of start and end points. Each instrument showed a scaled diagram of the actual block in color. The crossing options,

including both the path and the letter label, were coded in colors that were consistent across all instruments. For a given block, the start and end points for one of the two combinations were coded in red and the others in blue. Figure 1 shows an example of the instrument with a start-end combination in blue. Note that the duration of three minutes was chosen so that the participants can observe the street environment for a full signal cycle in most cases. Also the exact order of the options in an instrument depends on where the start and end points are located.

## **DATA COLLECTION**

Several aspects of the data collection logistics were discussed earlier. This section focuses on collection of static data and field surveys.

### **Static Data**

Data describing the static aspects of the street environment were collected while the survey instruments were being developed. A form was developed for field collection. It had a section for data related to crossing conditions at each of the five potential start points and a section for data related to the roadside environment. Before any data were recorded, block length was measured and each of the five possible start and end points were marked. In addition, the pre-selected combinations of start and end points were color-coded into blue or red as designed.

### **Field Surveys**

The final sample of 86 survey participants was recruited through a temporary staffing agency. The initial target sample size was 96 so that a total of 24 would participate on each of the four survey days with 12 on each bus. Ten did not show up for all four days combined. This approach to selecting participants gave greater certainty in the number of recruited participants who actually showed up. Given the fact that completing the field surveys for any given participant took about 5 hours, recruiting volunteers through random sampling of residents in the study area would not have worked as well.

Field surveys were conducted toward the end of April 2002. Prior to departing a central location each day, participants were given verbal instructions and a participant identification number (PIN) at random. The PINs were numbered consecutively from 1. After the briefing, those participants with even PINs boarded one bus and the others the other bus.

At each block, the participants from the same bus were divided into two groups of around 5 to 6 in each group. Two survey workers brought one group to the blue start point and the other to the red point. These two survey workers were supervisors for the two start points. Both of them recorded the PINs of those participants at their start point. Both were also responsible for distributing and receiving the instruments and checking whether the instruments were filled properly. One of them was a timer as well who not only determined when to start and end a particular crossing scenario but also recorded times. In addition, the timer had a sheet with all six blocks that color-coded the locations of the two start and end combinations and mid-block crosswalk marking. At the same time, one survey worker brought one red flag and one blue flag to mark the end points for the two combinations. Another brought the orange traffic cones to the appropriate locations if a mid-block crosswalk was required. Two survey workers got into position for collecting turning movements at intersections and two others for turning movements at mid-block locations (including driveway volumes and u-turns) as well as large vehicles (i.e., trucks, buses, and vans that are larger than regular household vehicles). These were the data collectors, who used pre-developed forms for these dynamic data.

Once everyone was in position, the timer signaled to everyone when to start a crossing scenario. Once started, the participants were given three minutes to observe the street environment as indicated on the survey instrument and were asked to fill the instrument right after being instructed to stop observation. Meantime, the data collectors were recording turning movements and the number of large vehicles. Also, pre-laid traffic counters were recording directional traffic volumes by speed ranges. Four two-way radios were used for communications. Once each group was done with its first crossing scenario for the block, the two groups then switched locations with each other. Once both scenarios were done for a given block, everyone boarded the bus and traveled to the next block.



## **Dataset**

A dataset was developed from the survey scenarios, static data, dynamic data, and stated preferences. It contained a total of 1,028 observations (out of 1,032 possible observations) and 42 independent variables. Among the variables, 6 are trip characteristics, traffic characteristics, roadway characteristics for crossing, and traffic control characteristics, respectively. Nine are roadside characteristics that are measured separately for each side of the blocks. The first two columns of Table 1 explain these variables.

### *Blocks*

The 48 blocks had a range of combinations of the potential determinants considered. The average length was 618 feet with a minimum of 232 feet, a maximum of 1,300 feet, and a standard deviation of 314 feet. There were 15 blocks on a 2-lane road, 16 on a 4-lane road, and 17 on a 6-lane road. Sixteen of these blocks were undivided; 20 were with restrictive medians (raised or grassy); and 12 were with painted medians. Crosswalk marking was present at both intersections for 7 blocks, at one intersection for 24 blocks, and at none of the intersections for 17 blocks.

### *Participants*

The sample of participants had more females than males but had a reasonable spread by age and household income. The 65+ age group crossed far fewer roads than the younger groups on the day before survey, ranging from one third of the average crossing by the 25-44 group to one half of the average by the other groups. On the day before survey, the female participants crossed roads one and half times versus three and quarter times by the male participants. Few of the participants perceived themselves having difficulty with walking at normal speed. These were evenly distributed between the two genders. Far more of them perceived having difficulty with walking at higher walking speeds, however, especially with the 45-64 group. Every one of the 13 participants who reported no difficulty at normal walking speed but reported difficulty at higher walking speeds was female.

### *Descriptive Statistics*

All potential independent variables were examined for correlation. For example, traffic volume is positively correlated with green time and crossing distance for each intersection option with a correlation coefficient slightly over 0.5. This information was then used later in model estimation. In addition, individual independent variables were examined for the reasonableness of their mean, standard deviation, maximum value, and minimum value.

## **ESTIMATION**

### **Hypotheses**

Hypotheses were formulated for a statistical model and expected directions of effects of the independent variables.

### *Statistical Model*

It was hypothesized that the most appropriate statistical model is the nested logit model (31). It is natural to view the six potential options for street crossing as two distinctive groups: those related to cross at intersections and those related to crossing at mid-block locations. That is, the nested logit model has a two-level structure. The top level has two branches: intersections (I) and mid-block locations (M). The bottom level has two options in the intersection branch (A and E) and up to four options in the mid-block branch (B, C, D, F).

### *Independent Variables*

The hypothesized direction of effects of independent variables was based on a basic specification of the utility functions. This specification involved two aspects. First, all variables were to be entered linearly to reduce the complexity of the model. Second, the specific utility functions to which a particular independent variable may enter were determined. Several criteria were used for this purpose. One criterion was whether an independent variable is constant across the options (e.g., roadside walking varies but not total traffic volume). One criterion was whether an independent variable is defined for each crossing option (e.g., signalization is defined for intersections only). Another criterion was whether a specific direction of effects could be hypothesized (The width of shoulders or bike lanes is likely to increase the probability that pedestrians choose options that require roadside walking but is likely to decrease the probability that they choose options that do not require such walking). Based on this specification, hypotheses were formulated for each independent variable. Table 1 also shows the specification and hypotheses for individual options and the two branches.

Table 1. Variables and Hypotheses

Variables		Hypotheses							
		Individual Options						Branches	
Description	Unit	A	E	B	C	D	F	I	M
<i>Trip</i>									
Walking distance	Feet along roadsides	-	-	-	-	-	-		
Crossing distance	Feet on travel lanes	-	-	-	-	-	-		
Start and end at mid-block locations	1 if true; 0 otherwise								+
Start at mid-block & end at intersection	1 if true; 0 otherwise								+
Start at intersection & end at mid-block	1 if true; 0 otherwise							+	
Start and end at intersections	1 if true; 0 otherwise							+	
<i>Traffic</i>									
Traffic volume	Vehicles per hour			-	-	-	-		
Mid-block running speed	Miles per hour			-	-	-	-		
Vehicle mix	Percent trucks			-	-	-	-		
Driveway volumes	Vehicles per hour			-	-	-	-		
Mid-block U-turns	Vehicles per hour			-	-	-	-		
Intersection turnings	Vehicles per hour	-	-						
<i>Roadway- Crossing</i>									
Right-turn lane	1 if present; 0 otherwise	-	-						
Left-turn lane	1 if present; 0 otherwise	-	-						
Acceleration lane	1 if present; 0 otherwise	-	-						
Crosswalk marking	1 if marked; 0 otherwise	+	+	+	+	+	+		
Restrictive medians	Width in feet	+	+	+	+	+	+		
Non-restrictive medians	Width in feet	-	-	-	-	-	-		
<i>Roadway- Roadside</i>									
Driveway frequency (nearside)	Number			+	+				
Sidewalk (nearside)	1 if present; 0 otherwise			-	-				
Buffer (nearside)	1 if present; 0 otherwise			-	-				
Barriers in buffer (nearside)	1 if present; 0 otherwise			-	-				
Curbed roadside (nearside)	1 if curbed; 0 otherwise			-	-				
Width of outside lane (nearside)	Feet			-	-				
Width of shoulder / bike lane (nearside)	Feet			-	-				
Driveway frequency (farside)	Number				+	+			
Sidewalk (farside)	1 if present; 0 otherwise				-	-			
Buffer (farside)	1 if present; 0 otherwise				-	-			
Barriers in buffer (farside)	1 if present; 0 otherwise				-	-			
Curbed roadside (farside)	1 if curbed; 0 otherwise				-	-			
Width of outside lane (farside)	Feet				-	-			
Width of shoulder / bike lane (farside)	Feet				-	-			
<i>Control</i>									
Traffic signal	1 if present; 0 otherwise	+	+						
Signal cycle length	Seconds	-	-						
Signal spacing	Feet to next signal	-	-						
Pedestrian signal	1 if present; 0 otherwise	+	+						
Green time	Seconds	-	-						
Green ratio	Unit-less	-	-						

## Model Estimation

Model estimation was a complex process because of the large number of variables and multiple utility functions involved. Model estimation followed two stages and multiple steps.

The first stage resulted in a basic model that included only those characteristics that were explicitly shown in the instruments: traffic signals, pedestrian signals, crosswalks, relative crossing distance, relative roadside walking distance, and the location of the start and end points. These characteristics were highly significant and showed the hypothesized direction of effects in the basic model. This stage followed three steps: 1) Estimated a nested logit model of our initial specification; 2) Deleted variables one at a time that were significant but contradicted our hypothesis; and 3) Deleted variables one at a time that were consistent with our hypothesis but were insignificant.

One example of the variables that were significant but contradicted our hypotheses was driveway frequency for each roadside. As indicated in Table 1, it is reasonable to expect that people would be more likely to take options that do not require walking along a roadside that has higher driveway frequency. That is, the coefficients for driveway frequency should be positive as specified. However, they were consistently significant and negative. It is difficult to determine the exact reason for this contradiction. One possible explanation is that driveway frequency is positively correlated with block length. People are less likely to take mid-block options along longer blocks. When block length is not used as an independent variable, the coefficients of driveway frequency may reflect the effects of block length rather than its own effect.

The second stage resulted in our preferred model that included three additional variables: traffic volume, width of a shoulder or bike lane on the nearside, and width of a shoulder or bike lane on the farside. This stage took an opposite approach from the first stage. This was done by starting with the basic model from the first stage and adding one variable at a time that was not already in the basic model. This stage also involved making tradeoffs between certain variables. Signal cycle, for example, made traffic volume become insignificant when both were present although traffic volume worked alone. Since the presence of traffic signals was already in the basic model, it was decided to keep traffic volume rather than signal cycle.

Table 2 presents our preferred model. It contains 10 variables descriptive of the street environment. The model also includes several alternative-specific constants and two inclusive values for the two branches. Note that the columns of coefficients are not in the same order as the options. The coefficients for the two intersection options, A and E, are placed next to each other first. They are followed by the coefficients for the mid-block options. The same order is used in the discussion below. For ease of reference, the individual options are redefined at the bottom of the table. The t-statistics are reported in the parentheses below the coefficients.

The model is well behaved. First, all variables are significant and have the hypothesized direction of effects. Second, it fits the data well. The  $\rho^2$  adjusted for the number of variables is 0.452. In contrast, it is common to see an adjusted  $\rho^2$  below 0.3 in discrete choice models such as mode choice models. Third, the model is consistent with utility maximization (32-34). The scale parameter at the bottom level of the nested logit model was scaled to 1. The estimated coefficients of the inclusive values fall between 0 and 1. Third, the estimated coefficients of the inclusive values are significantly different from 1, indicating that the nested logit model fits the data better than the conditional logit.

One way to understand the model is to look at the implied elasticities, which measure how responsive the choice probabilities are to changes in continuous variables. The model has three of these: crossing distance, roadside walking distance, and traffic volume.

- With respect to crossing distance, the elasticity is -0.099 (A-left intersection), -0.117 (E-right intersection), -0.050 (B-cross first and walk later), -0.584 (C-jaywalk), -0.057 (D-walk first and cross later), and -0.025 (F-midblock crosswalk). None of the options is responsive to changes in crossing distance. Option C (jaywalk), however, is far more responsive than the other options. That is, pedestrians are far less likely to jaywalk than to take other options when crossing distance increases.

Table 2. Nested Logit Model of Pedestrian Street Crossing Behavior (t-statistics in parentheses)<sup>1</sup>

Variable	Definition	Coefficient							
		Individual Options						Branches	
		Intersections		Mid-block				Inter-sections	Mid-block
		A	E	B	C	D	F	I	M
Alternative-specific constant	1			2.2079 (7.34) <sup>2</sup>		1.7266 (5.55)	1.3875 (3.44)	2.2332 (4.20)	
Walking distance	Feet along roadsides	-0.0034 (-11.65)	-0.0034 (-11.65)	-0.0034 (-11.65)	-0.0034 (-11.65)	-0.0034 (-11.65)	-0.0034 (-11.65)		
Crossing distance	Feet on travel lanes	-0.0027 (-2.31)	-0.0027 (-2.31)	-0.0027 (-2.31)	-0.0027 (-2.31)	-0.0027 (-2.31)	-0.0027 (-2.31)		
Start and end at mid-block locations	1 if true; 0 otherwise								1.5722 (3.14)
Start at mid-block & end at intersection	1 if true; 0 otherwise								0.8415 (2.32)
Traffic volume	Vehicles per hour			-0.0003 (-1.77)	-0.0003 (-1.77)	-0.0003 (-1.77)	-0.0003 (-1.77)		
Crosswalk marking	1 if marked; 0 otherwise	1.0002 (4.30)	1.0002 (4.30)	0.7891 (4.02)	0.7891 (4.02)	0.7891 (4.02)	0.7891 (4.02)		
Width of nearside shoulder/bike lane	Feet if present; 0 otherwise			-0.0728 (-1.22)	-0.0728 (-1.22)				
Width of farside shoulder/bike lane	Feet if present; 0 otherwise				-0.0923 (-1.42)	-0.0923 (-1.42)			
Traffic signal	1 if present; 0 otherwise	0.7502 (3.42)	0.7502 (3.42)						
Pedestrian signal	1 if present; 0 otherwise	1.2350 (4.34)	1.2350 (4.34)						
Inclusive value: Intersections	$J_I \equiv \ln(e^{U_A} + e^{U_E})$							0.7585 (7.05)	
Include value: Mid-block	$J_M \equiv \ln(e^{U_B} + e^{U_C} + e^{U_D} + e^{U_F})$								0.8342 (5.87)
Utility function	$\sum (\text{Variable} * \text{Coefficient})$	$U_A$	$U_E$	$U_B$	$U_C$	$U_D$	$U_F$	$U_I$	$U_M$
Number of Observations									1,028
Number Cases									4,334
Log likelihood with constants only									-1769.605
Log likelihood at convergence									-963.728
Unadjusted $\rho^2$									0.455
Adjusted $\rho^2$									0.453

<sup>1</sup> NLOGIT 3.0 of Econometric Software, Inc. was used to estimate this model with full information maximum likelihood. The RU1 normalization was used for the scale parameters. The nested logit model has two levels with variable options across observations. The top level has two branches: intersections and mid-block locations. The bottom level has two options in the intersection branch (A and E) and up to four options in the mid-block branch (B, C, D, F). A = Crossing at the left intersection (left intersection); B = Crossing at a mid-block start point at a right angle (cross first and walk later); C = Crossing with a jaywalk between the start and end points (jaywalk); D = Walking to the opposite of a mid-block end point and crossing there at a right angle (walk first and cross later); E = Crossing at the right intersection (right intersection); and F = Crossing at a mid-block crosswalk (mid-block crosswalk). I = Intersections; M = Mid-block. Left and right are determined in terms of the nearside. The nearside of a block is where the start point is.

<sup>2</sup> It is appropriate to determine the significance of the coefficients with a one-sided test because the null hypothesis for each coefficient is either being positive or negative rather than zero. A coefficient would be significant at the 10 percent, 5 percent, and 1 percent level if its t-statistic is at least 1.282, 1.645, and 2.326, respectively. These reported t-statistics do not correct for potential overestimation due to the repeated observations from individual respondents.

- With respect to roadside walking, the elasticity is -1.547 (A-left intersection), -1.853 (E-right intersection), -0.243 (B-cross first and walk later), -0.345 (D-walk first and cross later), and -0.232 (F-midblock crosswalk). The probability of an intersection being chosen is highly responsive. An increase of 10 percent in roadside walking could reduce the probability by 15 to 18 percent. In contrast, the probability of any mid-block option being chosen is irresponsive.
- With respect to traffic volume, the elasticity is -0.197 (B-cross first and walk later), -0.273 (C-jaywalk), -0.134 (D-walk first and cross later), and -0.059 (F-midblock crosswalk). Pedestrians are less likely to choose mid-block options when traffic volume increases. This impact, however, is irresponsive. Furthermore, the elasticity values for options B (cross first and walk later), D (walk first and cross later), and F (mid-block crosswalk) are several times higher in magnitude than those with respect to crossing distance but lower in magnitude than those related to roadside walking distance. For option C (jaywalk), however, the elasticity with respect to traffic volume is only half of that in magnitude as crossing distance.

To present the formula for probability calculations, let  $U_O$  ( $O = A, E; B, C, D, F; I, M$ ) be the sum of the products of all variables in the first column with the corresponding parameter values for option  $O$  on the right side columns in Table 1. Note that the inclusive values are  $V_I = \text{Ln}(e^{U_A} + e^{U_E})$  and  $V_M = \text{Ln}(e^{U_B} + e^{U_C} + e^{U_D} + e^{U_F})$  for the intersection and mid-block branches, respectively. The probability of a crossing option being chosen is the product of its marginal and conditional choice probabilities. The conditional probability represents the probability of choosing a particular crossing option once the choice has been made between intersections or mid-block options. With intersections being chosen ( $I$ ), for example, the conditional probability of intersection  $k$  ( $k = L, R$ ) being chosen is given by  $P(k / I) = e^{U_k} / e^{V_I}$ . With mid-block options being chosen ( $M$ ), similarly, the probability of mid-block option  $m$  ( $m = B, C, D, F$ ) being chosen is given by  $P(m / M) = e^{U_m} / e^{V_M}$ . The marginal probability represents the probability of choosing intersections or mid-block options. Specifically, the probability of either being chosen ( $J = I, M$ ) is  $P(J) = e^{U_J} / e^V$  where  $V = \text{Ln}(e^{U_I} + e^{U_M})$ .

## DISCUSSION

### *Limitations*

Before discussing potential implications, it is critical to understand the simplifications made as part of the research. One simplification is that the dynamics of traffic conditions and pedestrian's street crossing behavior are modeled away. The model relates the average traffic conditions during a three-minute period with how a pedestrian may have chosen to cross a street block under such average conditions. Whether safe traffic gaps are available can change quickly over time and across locations along a street block. Such temporal and spatial dynamics in traffic conditions lead to dynamics in the street crossing behavior of pedestrians as well. This simplification falls short for understanding certain crossing behavior, such as mid-block dash, i.e., situations where the pedestrian unexpectedly appeared in front of a motorist while the pedestrian was running and the motorist's view was not obstructed (35). Another simplification is that it ignores the role of time constraints. Relative to other direct attributes, time and its predictability would become far more important to a pedestrian when he has a tight time constraint. As a result, he may take riskier crossing options. By excluding time constraints, the usefulness of the model is reduced in understanding the behavior of transit users in trying to catch a coming bus on the other side of the road. The exclusion is made partly because of the difficulty in modeling time constraints.

### *Implications*

Implications relating to research, planning tools, and engineering solutions are discussed.

**Research Methods.** A number of implications can be drawn that have both current and lasting value to researchers. These include:

1. The results show that pedestrian street-crossing behavior can be reasonably modeled with indirect factors that can be directly measured in practice. In this case, the indirect factors describe the street environment. However, an otherwise similar model based on direct factors alone fits the reported pedestrian street-crossing behavior better. In fact, the adjusted  $p^2$  increased from 0.453 to 0.552. The direct factors measure perceived safety, time, and predictability on a scale from 1 (least favorable) to 10 (most favorable). The data were collected from the respondents in the field just after they stated their crossing preference for each crossing scenario.

2. Excluding personal attributes from the preferred model appears to have small impacts on the model. An alternative model with added personal attributes was estimated. The addition improved the preferred model with an increase in the adjusted  $\rho^2$  to 0.471. The elasticity with respect to roadside walking was compared, for example, and it increased from  $-1.547$  to  $-1.593$  for the left intersection and from  $-1.853$  to  $-1.901$  for the right intersection.
3. The reported results earlier show that the nested logit model fits the stated pedestrian street-crossing behavior better than the conditional logit model.
4. The quasi-stated preference approach provides an alternative to the standard stated-preference approach.
5. The survey design provides an example of modeling the continuum of street crossing options in real life with discrete methods.

**Planning Tools.** The existing tools for determining pedestrian level of service are based on simple regression models that predict pedestrian perceptions of quality of service with the street environment. The estimated model from this research could provide a new approach that is based on pedestrians' overall satisfaction with street crossing. Specifically, the estimated utility functions can be combined to provide a meaningful measure of the overall satisfaction from crossing specific blocks:  $V = \ln(e^{U_l} + e^{U_m})$ . This concept is similar to using the denominator of a logit destination choice model as an accessibility measure (36). More important, this new approach to determining pedestrian level of service is also a behaviorally sound way to measure level of service across different modes equally. The National Corporate Highway Research Program has planned a research project to look for a unified approach for equal measurement of level of service across modes (37).

**Engineering Solutions.** The estimated model may be used to simulate how certain engineering solutions may influence how pedestrians cross streets.

1. The model can be used to determine the circumstances under which pedestrians are more likely to go to an intersection or a mid-block crosswalk. With some basic assumptions, curves may be developed to show how different combinations of selected aspects of the street environment influence the likelihood that a typical pedestrian would choose an intersection or a mid-block crosswalk in daytime conditions.
2. The model can also be used to determine how marking a mid-block crosswalk may discourage pedestrians from taking risky options.
3. Transit stops are often the destination of pedestrians crossing a street. When these stops are located inappropriately, transit users may be more likely to take risky options for crossing. For given origins, the model can help understand how the destination within a block can influence the likelihood of pedestrians to take risky options. The same implication also applies to locating walkways from major activity centers, newspaper boxes, vending machines, etc.

The actual simulation requires additional space to explore and may be carried out in a later paper.

## ACKNOWLEDGEMENTS

The Florida Department of Transportation provided funding through the National Center for Transit Research. We like to thank District Seven of the Department for performing the traffic counts and the HARTline for providing bus transportation.

## REFERENCES

1. Rouphail, Nagu M., Joseph E. Hummer, Joseph S. Milazzo II, and D. Patrick Allen (1998), Recommended Procedures for Chapter 13, "Pedestrians," of the HCM, FHWA-RD-98-107.
2. Hidas, Peter, Kolita Weerasekera, and Michael Dunne (1998), Negative Effects of Mid-block Speed Control Devices and their Importance in the Overall Impact of Traffic Calming on the Environment, Transportation Research-D 3, pp. 41-50.
3. Hunt, John, and Jalal Abduljabbar (1993), Crossing the Road: A Method of Assessing Pedestrian Crossing Difficulty, Traffic Engineering and Control 34, pp. 526-532.
4. Knoblauch, Richard L., Martin T. Pietrucha, and Marsha Nitzburg (1996), Field Studies of Pedestrian Walking Speed and Start-up Time, Transportation Research Record 1538, pp. 27-38.
5. Oppenlander, Joseph C. (1996), Correlation of Volumes with Acceptable Gaps in Pedestrian Signal Warrant, in Compendium of Technical Papers for the 66<sup>th</sup> ITE Annual Meeting, pp. 158-162.

6. Connelly, Marie L., Helen M. Conaglen, Barry S. Parsonson, and Robert B. Isler (1998), Child Pedestrians' Crossing Gap Thresholds, *Accident Analysis and Prevention* 30, pp. 443-453.
7. Milazzo II, Joseph S., Nagui M. Roupail, Joseph E. Hummer, and D. Patrick Allen (1998), Quality of Service for Interrupted Pedestrian Facilities in the 2000 Highway Capacity Manual, TRB Paper No. 99-0131.
8. Sisiopiku, Virginia P., and Darcin Akin (2000), Assessment of Pedestrian Crossing Options, Transportation Research Board Paper No. 00-1494.
9. Sharples, J.M., and J.P. Fletcher (2001), Pedestrian Perceptions of Road Crossing Facilities, Transport Research Laboratory, the Scottish Executive Central Research Unit, Edinburgh.
10. Baltes, Michael R., and Xuehao Chu (forthcoming), Pedestrian Level of Service for Mid-block Street Crossing, Transportation Research Record.
11. Crider, Linda B., Jodi Burden, and Feng Han (2001), Multimodal LOS -- Point Level of Service Project -- Final Report, Department of Urban and Regional Planning, University of Florida, Gainesville, Florida.
12. Sarkar, Sheila (1993), Determination of Service Levels for Pedestrians, with European Examples, Transportation Research Record 1405, pp. 35-42.
13. Sarkar, Sheila (1995), Evaluation of Safety for Pedestrians at Macro- and Microlevels in Urban Areas, Transportation Research Record 1502, pp. 105-118.
14. Kochevar, Robert A., and Nazir Lalani (1985), How Long Should a Safe Pedestrian Clearance Interval Be? ITE Journal 55, pp. 30-32 and 49.
15. Lalani, N. (2001), Alternative Treatments for At-Grade Pedestrian Crossings, Institute of Transportation Engineers, Washington, D.C.
16. Zegeer, C.V., J.R. Stewart, H Huang, and P. Lagerwey (2001), Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations: Analysis of Pedestrian Crashes in 30 Cities, Transportation Research Record 1773, pp. 56-68.
17. Van Houten, R., JEL Malenfant, and D. McCusker (2001), Advance Yield Markings: Reducing Motor Vehicle-Pedestrian Conflicts at Multilane Crosswalks with Uncontrolled Approach, Transportation Research Record 1773, pp. 69-74.
18. Huang, Herman F., and Michael J. Cynecki (2000), Effects of Traffic Calming Measures on Pedestrian and Motorist Behavior, Transportation Research Record 1705, pp. 26-31.
19. Schnell, T., F. Aktan, D.V. McGehee, M. Dvorak, J. Hunt, A. Reyes, and D. Sorak (2001), Pedestrian Visibility under Automobile Low-Beam Headlight Illumination: With and Without Headlight Covers, Transportation Research Record 1773, pp. 114-126.
20. Raymond, Paula D., and Richard L. Knoblauch (2000), the Effects of Crosswalk Markings on Vehicle Speeds, TRB Paper No. 00-1602.
21. Gordon, Stewart, and H. Douglas Robertson (1988), A Study of Driver Noncompliance with Traffic Signals, Transportation Research Record 1168.
22. Grayson, G. B. (1987), Pedestrian Risk in Crossing Roads: West London Revisited, Traffic Engineering and Control 28, pp. 27-30.
23. Thompson, Larry A. (1996), Pedestrian Road Crossing Safety, CPL Bibliography 331/332/333, Journal of Planning Literature 11, pp. 263-300.
24. Palamarthy, Srinivas, Hani S. Mahmassani, and Randy B. Machemehl (1994), Models of Pedestrian Crossing Behavior at Signalized Intersections, Research Report 1296-1, Center for Transportation Research, University of Texas at Austin.
25. Hamed, Mohammad M (2001), Modeling Pedestrians' Crossing Behavior: Some Empirical Evidence, in Road Safety on Three Continents, Proceedings of the Conference, 19-21 September 2001, Moscow, Russia.
26. Himanen, V., and R. Kulmala (1988), An Application of Logit Models in Analyzing the Behavior of Pedestrians and Car Drivers on Pedestrian Crossings, *Accident Analysis and Prevention* 2, pp.187-197.
27. Chu, Xuehao (2002), A Nested Logit Model of Pedestrian Street-Crossing Behavior, prepared for presentation at the 49<sup>th</sup> Annual North American Meetings of the Regional Science Association International, San Juan, Puerto Rico, November 13-16, 2002.
28. Chu, Xuehao (2002), Alternative Preference Methods, prepared for possible presentation at the 10<sup>th</sup> International Conference on Travel Behavior Research, Lucerne, August 10-15, 2003.
29. Chu, Xuehao, and Michael R. Baltes (forthcoming), Measuring Pedestrian Level of Service for Mid-block Street Crossing: The Selection of Potential Determinants, Transportation Research Record.
30. Landis, Bruce W., Venkat R. Vattikuti, Russell M. Ottenberg, Douglas S. McLeod, and Martin Guttenplan (2001), Modeling the Roadside Walking Environment: A Pedestrian Level of Service, TRB Paper No. 01-0511.
31. Greene, William H. (1990). *Econometric Analysis*, Macmillan Publishing Company, New York, NY.

32. Koppelman, Frank S., and Chieh-Hua Wen (1998), Alternative Nested Logit Models: Structure, Properties and Estimation, *Transportation Research B*, 32, pp. 289-298.
33. Hunt, Gary L. (2000), Alternative Nested Logit Model Structures and the Special Case of Partial Degeneracy, *Journal of Regional Science*, 40, pp. 89-113.
34. Hensher, David A., and William H. Greene (2002), Specification and Estimation of the Nested Logit Model: Alternative Normalizations, 36, pp. 1-17.
35. National Center for Bicycling and Walking (2002), Education and Public Awareness Campaigns, [http://www.bikefed.org/education\\_and\\_public\\_awareness\\_campaigns.htm](http://www.bikefed.org/education_and_public_awareness_campaigns.htm).
36. Ben-Akiva, Moshe, and Steven R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, the MIT Press, Cambridge, Massachusetts.
37. National Cooperative Highway Research Program (2002), Multimodal Arterial Level of Service, Project 3-70, FY 2003, <http://www4.trb.org/trb/crp.nsf/NCHRP+projects>.